2010-09-24

Systems Engineering Applied Leading Indicators: Enabling Assessment of Acquisition Technical Performance

Paul Montgomery

NPS-AM-10-175

ACQUISITION RESEARCH SPONSORED REPORT SERIES

Systems Engineering Applied Leading Indicators: Enabling Assessment of Acquisition Technical Performance

24 September 2010

by

Paul Montgomery, Associate Professor

Ron Carlson, Professor of Practice

Graduate School of Engineering & Applied Sciences

Naval Postgraduate School

Approved for public release, distribution is unlimited.

Prepared for: Naval Postgraduate School, Monterey, California 93943

The research presented in this report was supported by the Acquisition Chair of the Graduate School of Business & Public Policy at the Naval Postgraduate School.

To request Defense Acquisition Research or to become a research sponsor, please contact:

NPS Acquisition Research Program
Attn: James B. Greene, RADM, USN, (Ret)
Acquisition Chair
Graduate School of Business and Public Policy
Naval Postgraduate School
555 Dyer Road, Room 332
Monterey, CA 93943-5103
Tel: (831) 656-2092
Fax: (831) 656-2253
e-mail: jbgreene@nps.edu

Copies of the Acquisition Sponsored Research Reports may be printed from our website www.acquisitionresearch.org

Abstract

This paper discusses research in developing DoD acquisition metrics associated with Systems Engineering activities that may provide greater insight into the technical performance of development programs. These metrics are called Systems Engineering Applied Leading Indicators (ALI). We examine current development of single- and multi-factor ALIs that have been developed during the past year at the Naval Air Systems Command (NAVAIR) in Patuxent River, MD. The development methods, early examination of ALI utility, and user acceptance are discussed. The authors have been embedded with the NAVAIR Systems Engineering Development and Implementation Center (SEDIC) (the center of this work for NAVAIR) as part of this ALI exploration.

Keywords: DoD acquisition metrics, Systems Engineering Applied Leading Indicators (ALI), single- and multi-factor ALIs

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

The authors have been working in close collaboration with the members of the NAVAIR SEDIC team at Naval Air Station, Patuxent River, MD in the conduct of the subject research. Much of the foundational work referred to in this paper is a result of their efforts. The team includes Juan Ortiz, John Quartuccio, Paul Hood, and Javier Sierra of NAVAIR, and Gregory Hein, Johnathan Gilliard, and Tina Denq from Booz Allen Hamilton.

THIS PAGE INTENTIONALLY LEFT BLANK

About the Authors

Dr. Paul Montgomery: After retiring in 1990 from a twenty-year career in the Navy, Dr. Montgomery served as a Senior Systems Engineer with Raytheon and Northrop Grumman corporations and developed communications, surveillance, and sensor systems for commercial, military (USN, USA, USAF), and intelligence (NSA, NRO) communities. He earned his doctorate in systems engineering from George Washington University (D.Sc. '07) performing research related to cognitive/adaptive sensors, his MSEE (1987) from the Naval Postgraduate School, and his BSEE (1978) from Auburn University.

Dr. Paul Montgomery, Associate Professor
Graduate School of Engineering & Applied Sciences
Naval Postgraduate School
Monterey, CA 93943-5000
Tel: 703-568-1165
E-mail: prmontgo@nps.edu

Ron Carlson: Ron served twenty-six years in the US Navy as a pilot. During his last seven years on active duty, he led NAVAIR systems engineers through several years of systems engineering revitalization. Subsequent to his retirement from the Navy, he joined the NPS SE Department as a Professor of Practice. He is currently in the systems engineering doctoral program at Stevens Institute of Technology. He earned master's degrees in strategic studies and national policy from the Naval War College and Business Administration from Embry Riddle Aeronautical University, and a bachelor of science in nuclear engineering from the University of Michigan.

Ron Carlson, Professor of Practice
Graduate School of Engineering & Applied Sciences
Naval Postgraduate School
Monterey, CA 93943-5000
Tel: 301-757-9743
E-mail: rrcarlson@nps.edu

Both Dr. Montgomery and Ron Carlson are embedded faculty members in the SE Department who provide onsite research and instruction support to NAVAIR (Patuxent River, MD), NAVSEA (Dahlgren, VA, and Carderock, MD), and other East Coast NPS SE students.

NPS-AM-10-175

ACQUISITION RESEARCH
SPONSORED REPORT SERIES

Systems Engineering Applied Leading Indicators: Enabling Assessment of Acquisition Technical Performance

24 September 2010

by

Paul Montgomery, Associate Professor

Ron Carlson, Professor of Practice

Graduate School of Engineering & Applied Sciences

Naval Postgraduate School

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

THIS PAGE INTENTIONALLY LEFT BLANK

1. Introduction and Problem Definition

Background

What is the role of systems engineering (SE) in the acquisition and development of systems? The professional society for SE (INCOSE) defines SE as follows:

> Systems engineering is an interdisciplinary approach and means to enable the realization of successful systems. It focuses on defining customer needs and required functionality early in the development cycle, documenting requirements, and then proceeding with design synthesis and system validation while considering the complete problem: operations, cost and schedule, performance, training and support, test, manufacturing, and disposal. SE considers both the business and the technical needs of all customers with the goal of providing a quality product that meets the user needs. (INCOSE, 2010)

The principles, practices, and methods of SE are well defined and long practiced by Government and industry (INCOSE, 2010; NASA, 2007; Secretary of the Navy, 2008). The value added by disciplining the development of a system is well appreciated and in the mid 1990s, SE practices were augmented with the concepts of SE metrics (INCOSE, 1995, 1998; Roedler, 2005). Early implementation of these metrics has been directed at measuring the performance of the SE process itself.

In the Weapons Systems Acquisition Reform Act (WSARA, 2009), systems engineering authorities, practices, and imperatives are reemphasized throughout. (Systems engineering is mentioned 45 times.) New requirements exist for performance assessment and root cause analysis that will require insights into engineering metrics, some of which could include the leading indicators discussed in this paper.

A special emphasis of the above SE definition is the consideration of not only the development team, but also all customers and stakeholders who are maximally interested in a project/program that is delivered satisfying cost, schedule, as well as technical goals. There is now interest within the SE community (Rhodes, Valerdi, & Roedler, 2009) on how to expand, define, and derive metrics and methods that would provide predictive or prognostic indicators of the success of a development effort as a whole (see Figure 1). While the existing SE metrics and methods have typically produced lagging and inferred indicators of the health and status of a development effort, current efforts and research are now underway to examine how to provide direct *leading indicators*, derived from SE and *applied* to understanding and predicting the technical trajectory of the aggregate development effort. Because we are *applying* and focusing the concepts of SE leading indicators (Roedler, Rhodes, Schimmoller, & Jones, 2010), we will refer to this concept as SE *Applied* Leading Indicators (ALI) for the remainder of this paper.

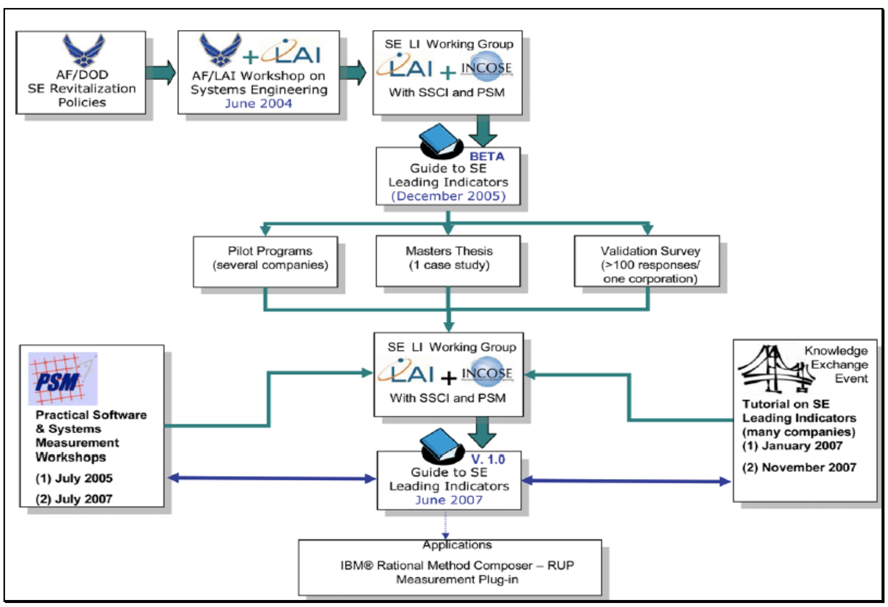


Figure 1. Government/Industry Partnership Exploring SE Leading Indicator Concepts and Application
(Roedler & Rhodes, 2007)

The authors set out attempting to focus on why programs fail to meet user expectations at delivery. Our goal was to determine what engineering metrics could be defined and analyzed to provide such insight where programs are apparently not getting such insight today (based upon failure rates of system qualification testing results). This goal led us to intersect ongoing efforts related to SE ALIs that we determined would provide an understanding of closely related metrics and processes that would underpin our investigation. The authors have been supporting and co-researching with Naval Air Systems Command (NAVAIR) in Patuxent River, MD, to examine the identification, relevance, and application of SE ALIs. NAVAIR has been examining the ALI concept through engagement with acquisition offices, data gathering and analysis, formulation of predictor algorithms, and prototype ALI tool development. The Systems Engineering Development and Implementation Center (SEDIC) is conducting this NAVAIR effort in collaboration with working groups depicted in Figure 1.

Problem Definition

Program managers apply well-proven and refined program metrics and control mechanisms largely based upon Earned Value Management (EVM). The EVM cornerstone metrics are cost and schedule each of which reference analysis of variances from plans and estimates. From EVM analysis, program cost and schedule status can be assessed and projection of those parameters can be inferred. Program managers, however, are not provided abundant metrics that can provide insights into the technical health of a development effort and indications of the trajectory of program health, good or bad. Risk metrics and processes provide some indications of technical health but are often qualitative and provide little algorithmic opportunities for prognostics. In general, program managers are faced with the development of complex systems, and they use EVM and risk management effectively; however, programs are failing to fully control costs and can routinely exceed cost estimates by 25% or more (see Figure 2).

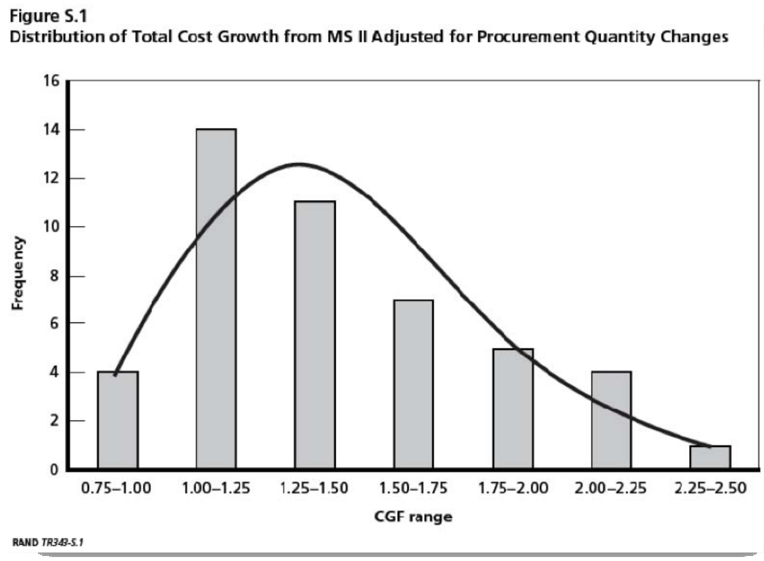


Figure 2. Control of Cost Growth of Programs Remains a Challenge
(Arena, Robert, Murray, & Younossi, 2006)

In addition to the quantity of programs that exceed cost estimates, it appears that acquisition cost growth can be attributed to causes centered upon control of technical baselines (see Figure 3). The development of ALIs is intended to gain much more granular insight into the development of the technical baselines as soon as possible to allow for both assessment and predicted program performance so mitigation can be applied. In summary, the specific problem and research response follow:

Problem—Program managers do not have access to adequate technical metrics in order to provide high fidelity assessment of technical health of a complex system development program and quantitative prediction of technical performance.

Research Question—Can SE technical metrics be identified, quantified, and methodically applied to complex system developments to provide technical assessment and leading indications of technical program performance and ultimate success?

Research Objectives:

- Identify relevant data supporting the development of ALIs
- Identify leading indicators tailored to systems engineering effectiveness
- Prototype ALI user tools to measure relevance and acceptance, and to obtain feedback
- Identify new, revised, or derived metrics to support refined ALI methods

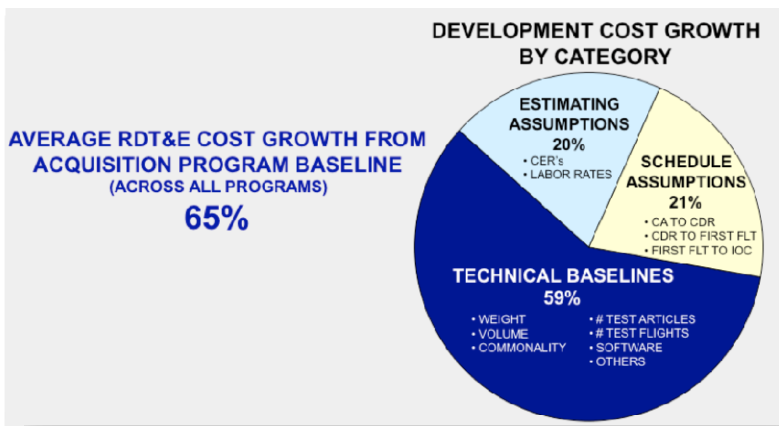


Figure 3. Cost Growth Largely Impacted by Control of Key Attributes of Technical Baselines
(Hein, 2009)

THIS PAGE INTENTIONALLY LEFT BLANK

2. Applied Leading Indicator Concepts

Technical Measurements

SE processes provide metrics, measurements, and analysis activities throughout systems development. These technical measurement activities provide insight into project technical performance and associated risks for lead system engineers and project managers. These metrics support larger top level measures including Measures of Effectiveness (MOEs), Measures of Performance (MOPs), Technical Performance Measures (TPMs), Key Performance Parameters (KPPs), and Key System Attributes (KSAs). These measures and metrics are qualified through continual testing and often manifest themselves graphically using control chart methods (see Figure 4).

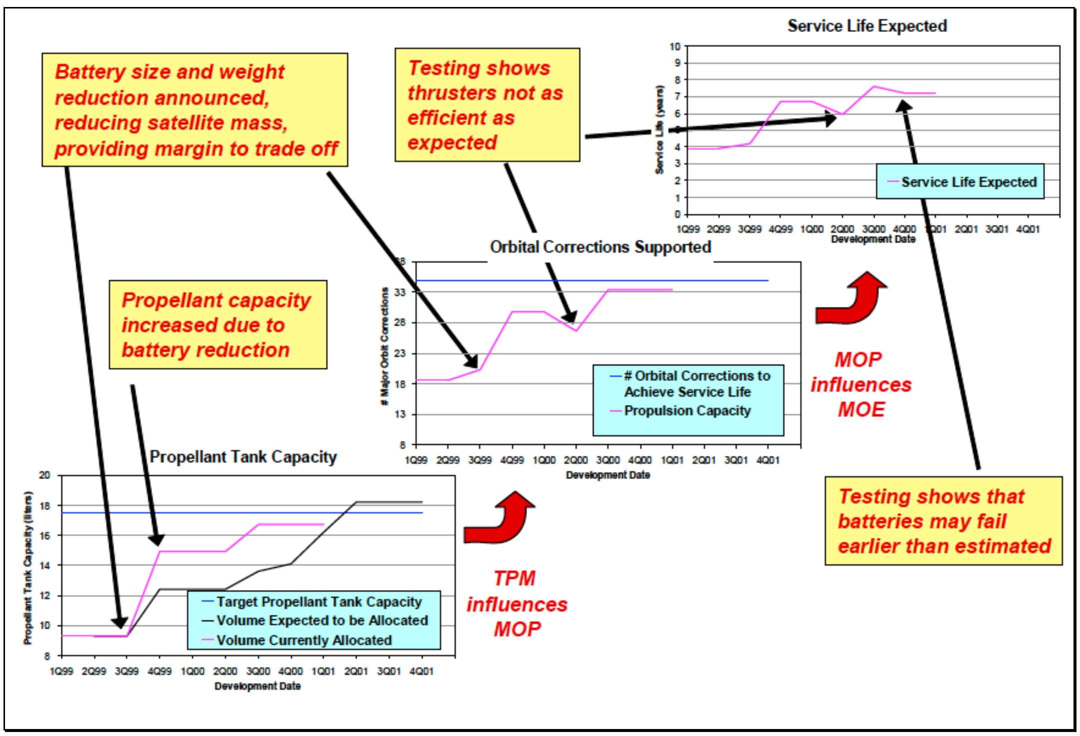


Figure 4. Technical Measures Associated with MOEs, MOPs, and TPMs Guide the Testing and Achievement of Specifications
(Roedler, 2005)

The above technical measurement processes are often focused on assessing the progress of the system in meeting specifications as development unfolds. Although the development of ALIs seems similar to these practices, the intent of ALIs is to provide a more holistic and prognostic assessment of the technical aspects of the project by integrating both system technical metrics as well as systems engineering-derived process metrics. ALIs, although substantiated in historical performance of similar projects, are highly forward-looking and technically-rich in fidelity. They are intended to inform the project technical approach and be fully integrated with the program management approach (see Figure 5).

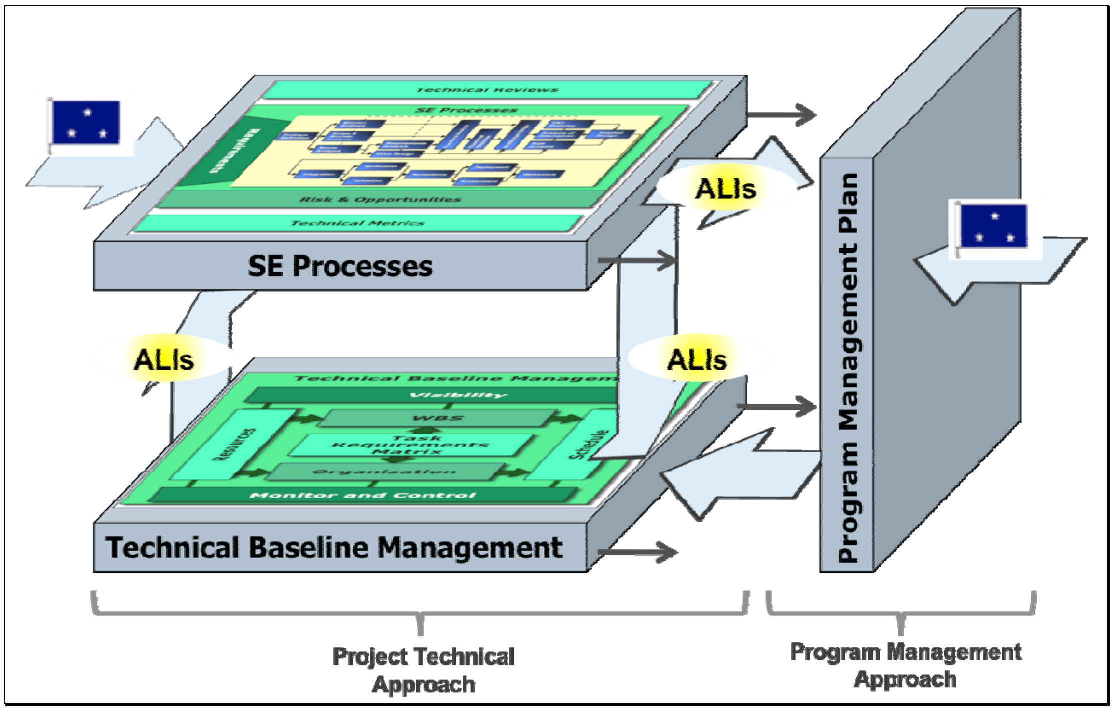


Figure 5. Alis Provide Metrics Rooted in SE Technical Approach and Supports Program Management Approach

The development and use of ALIs are intended to augment existing program/project management methods, not replace them. Although influenced by many similar metrics (e.g., cost, schedule, etc.), ALIs are derived from system attribute and system engineering metrics to produce technical health and prognostics that enhance the program manager’s overall assessment and direction of the project (see Figure 6). They enrich the existing EVM-derived assessment to provide project leadership higher fidelity project technical status and direction that enable greater decision analysis completeness.

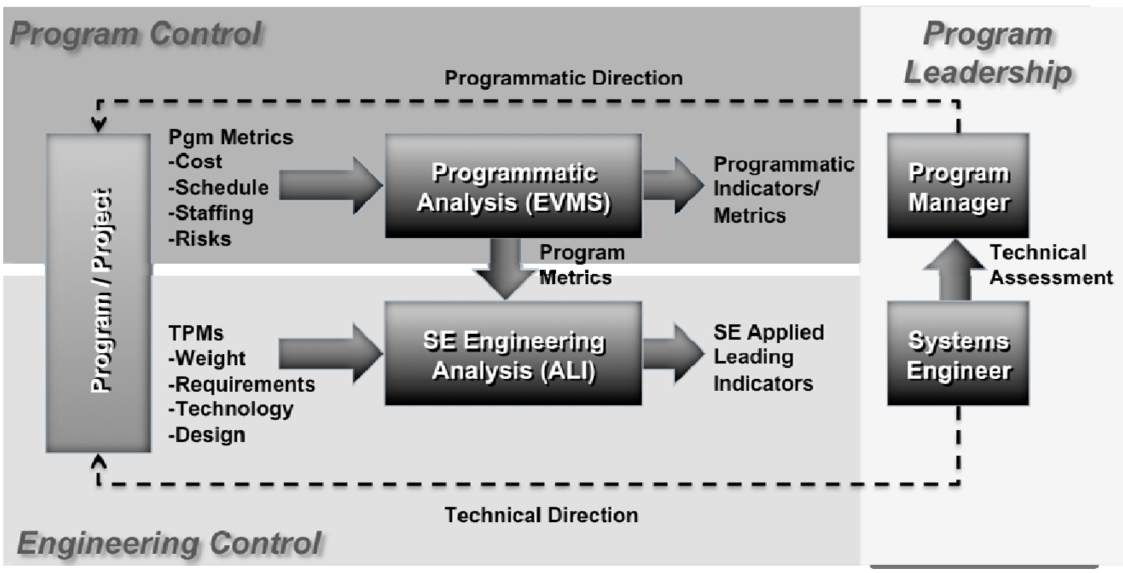


Figure 6. SE Applied Leading Indicators (ALI) Augment Program Management

THIS PAGE INTENTIONALLY LEFT BLANK

3. ALI Technical Approach

ALI Models and Tool Goals and Objectives

In support of the previously mentioned objectives, NAVAIR set out to integrate the technical resources and databases into an ALI methodology that can be integrated into NAVAIR acquisition business practices. The primary goals of the NAVAIR ALI effort are as follows:

- To find and assess data repositories of program data with sufficient content and relevance to support development of ALIs,
- To develop an understanding of the relationships between key technical factors and the performance of the acquisition program and,
- To develop models and tools to assist the acquisition management team to gain a greater insight into the technical performance of their program.

The first step of gathering data was, and continues to be, a challenge. Although NAVAIR has rich data repositories, several factors must be considered during collection to ensure relevance. Some factors include: availability of technical data with metrics, understanding of the metrics across different organizations, common taxonomy, accuracy of the metrics, sufficient breadth and depth, sufficient sample sizes for credible statistical analysis, reconciling different development cycle, etc. Examples of candidate technical metrics are the following:

- Aircraft empty weight,
- Software metrics,
- Architecture metrics,
- Requirements metrics,
- Closure rates of discrepancies from technical reviews
- Reliability, Availability, and Maintainability (RAM) metrics,
- Technical risk metrics,

- Engineering staffing metrics,
- System complexity, and
- Technology maturity.

During the early research efforts, aircraft weight was determined to be a prime candidate for investigation as a key technical metric. As discussed in Hess, and Romanoff (1987) and in Large, Campbell, and Cates (1976), the cost associated with the development of aircraft and their systems can be highly dependent on weight. This association was confirmed to hold true at NAVAIR as discussed in the next section. The NAVAIR Mass Properties Division has a rich database of weight status reports for most large NAVAIR programs and the NAVAIR Cost Department has monthly data for all major aircraft development contracts. As will be shown, we started with weight versus cost data as our first ALI to analyze.

The data was collected to form a historical baseline of program performance of similar or related programs. (Later ALI phases would incorporate current program data to predict future performance.) The data was also "affinitized" or grouped in like-program categories to maintain relevance of analysis results. Examples of these groupings included aircraft development with similar plan forms (e.g., rotary, fixed wing, remotely piloted, etc.), size of the program (ACAT I, II, etc.), and mission (fighter, transports, etc.). In all, approximately 11 programs form the foundation for data analysis. The following section details the method employed throughout this research and the development of ALI models and tools.

ALI Method

The ALI process objectives were to gain an understanding of the data, relationships, statistical saliencies, algorithms, and ultimately, the development of an ALI tool, which is shown in Figure 7. (For additional amplification of this approach, see Appendix A in Roedler, Rhodes, Schimmoller, & Jones, 2010). The overall process flow starts by determining key interactions among technical factors and the program performance, analyzing relationships, developing models and ALI tools, and seeking user inputs and feedback on the ALI tools. The analytical step performs statistical correlation, regression, or sensitivity analyses. The modeling and tool development is accomplished in Microsoft Excel using Visual Basic for Applications, which is the underlying programming language for Microsoft applications.

Data is drawn from NAVAIR data repositories as input to each process step. Users are engaged throughout this process for suggestions on data relevance, algorithm relevance, tool design, and tool utility.

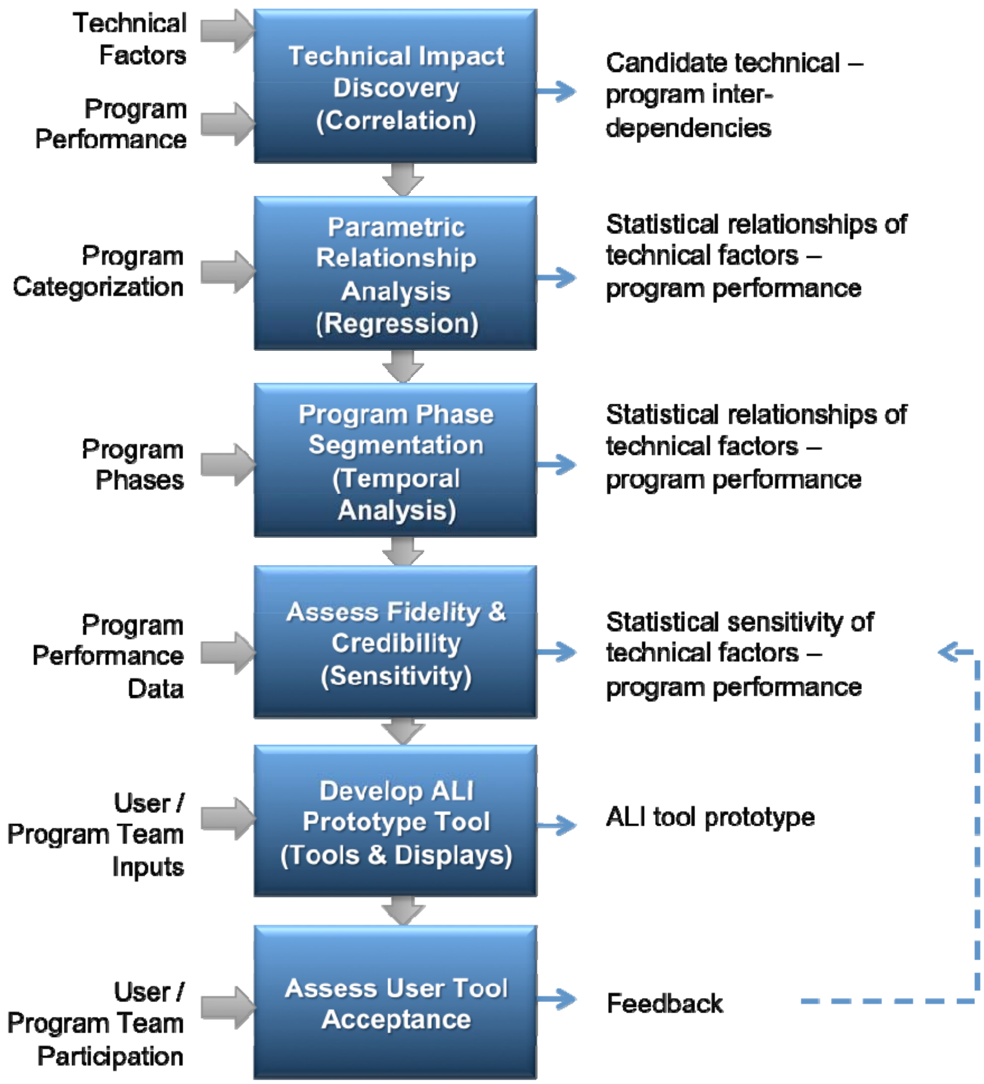


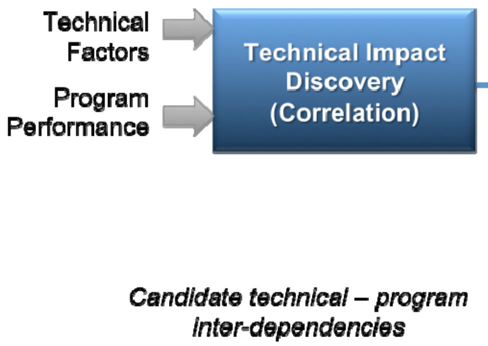
Figure 7. SE ALI (Single-Factor) Analysis, Modeling, and Prototype Tool Development Process

Although the Figure 7 process depicts the single-factor ALI analysis and modeling, the process is equally applicable for multi-factor analysis. The multi-factor approach perspective is discussed in subsequent sections.

Discover the most influential technical factors impacting program performance

As previously discussed in the section titled ALI Models and Tool Goals and Objectives, a variety of technical factors are candidates to be investigated to determine the impact to program performance. The first step in our process is to determine which of the technical factors have a key impact on the overriding program performance parameters, cost, and schedule.

As shown in the example in Figure 8, a correlation matrix is developed to correlate each technical factor metric against program performance measures (cost and schedule). In the example shown in Figure 8, aircraft weight is shown correlated against program performance, although several technical factors were examined prior to selecting aircraft weight as our first ALI. This correlation process identifies whether or not there is a significant influence on program performance from aircraft weight. Large positive correlation values in the cells of interest provide strong indication of the correlation relationships.



Technical Factor (Aircraft Weight) -to-Program Performance (Cost Growth) Correlation		Technical Factors: % Cumulative Weight Growth	Technical Factors: % Below Not-to-Exceed Weight	Technical Factors: % Below Weight Plan	Program Perf: Pgm Cost Performance Index	Program Perf: Pgm Schedule Performance Index	Program Perf: % Cost Growth
Technical Factors	% Cumulative Weight Growth	1.000	0.414	0.665	0.186	0.908	0.808
Technical Factors	% Below Not-to-Exceed Weight	0.414	1.000	0.812	0.457	0.949	0.084
Technical Factors	% Below Weight Plan	0.665	0.812	1.000	0.004	0.036	0.031
Program Perf	Pgm Cost Performance Index	0.186	0.457	0.004	1.000	0.179	0.085
Program Perf	Pgm Schedule Performance Index	0.908	0.949	0.036	0.179	1.000	0.320
Program Perf	% Cost Growth	0.808	0.084	0.031	0.085	0.320	1.000

=correlations of interest

Figure 8. Correlation of Technical Factors to Program Cost Growth Parameters Leads to Candidate ALI Factors

The correlation matrix shows the Pearson's *R* correlation coefficient for each parameter pair. Each technical versus performance parameter pair is tested for statistical significance by Student's *t* statistic.

$$t_{N-2,\alpha} = \frac{R_c}{\sqrt{(1-R_c^2)/(N-2)}} \qquad \text{(Equation 1)}$$

- *N* is the number of data points for a technical parameter versus performance parameter pair.
- $t_{N-2,\alpha}$ is the Student's *t* statistic for *N*–2 degrees of freedom.
- α = 0.05.
- Eliminate all parameter pairs where the coefficient of correlation is less than the critical value R_c.

It should be noted that the technical data is often not usable across multiple platforms to the same level of equivalence because different units are applied (e.g., pounds, kilograms, etc). This makes model aggregation problematic. Additionally, incongruent scale of aircraft also makes the use of absolute values illogical (e.g., an unmanned air system (UAS) is much smaller and lighter than a fighter aircraft). We, therefore, transformed absolute weight values into relative weights for our analyses. We related weights to percentages such as percent below weight plan (%BP), percent below not-to-exceed weight limit (%BNTE), and percent cumulative weight growth from original estimate (%CWG) for our analysis and modeling. Similarly, percentages were used for program performance metrics, especially percent cost growth (%CG).

Analyze statistical relationships and develop parametric models describing coupling among technical factors and program performance

From the previously described correlation analysis, candidate technical factors emerged that had significant influence on program performance. We selected aircraft weight as our first parameter. The next step in our analysis was to determine if the weight growth data has predictive strength in predicting cost growth. We examined this predictive strength through regression analysis. We employed linear regression because it proved to be as effective as the non-linear methods (exponential and polynomial) that we examined. Our regression analysis revealed significant statistical strength of using weight-growth as a cost-growth predictor across several programs (see Figure 9).

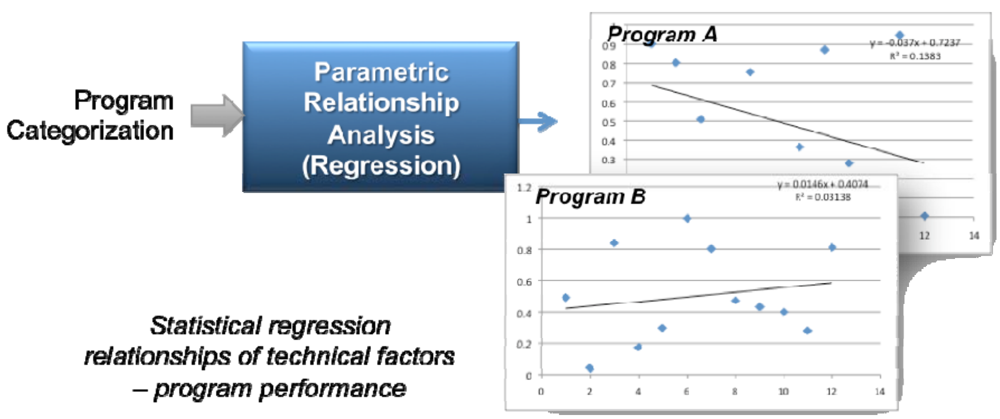


Figure 9. Regression Analysis Provides Basis for Algorithmic Description of ALI Factor Impacts on Program Performance

We examined the logic of the slopes, significance of intercepts, goodness of fit (R^2), randomness of residuals, and correlation relevance (compare fit, R^2, with R_c from the correlation process). The regression validity was compared interprogram and intraprogram, and we found that separation/affinitization of regression into closely related program categories was appropriate and necessary. Examples of categories used for NAVAIR aircraft programs included:

- Mission Type,
- Program Executive Office (PEO),
- Conventional Take Off and Landing (CTOL) versus Vertical Take-Off and Landing (VTOL), and
- Fixed wing versus rotary wing.

The regression statistics demonstrated discontinuities within program categorization. It was determined that, in addition to program affinitization of the regression analysis, additional time segmentation would be necessary. This segmentation is discussed in the next section.

Analyze impact of time and program phases on parametric relationships and models

The regression results showed significant statistical strength of using weight-growth as a cost-growth predictor; however, the data must be segmented into major epochs of program development to maximize this predictive strength. The epochs were separated by major design reviews (e.g., Preliminary Design Review (PDR), Critical Design Review (CDR), etc.) to ensure predictive usefulness. During this time segmentation process, we also aggregated the regression analysis for each program phase such that a single, significant predictor emerged for each phase. The result is a family of predictors of cost growth (based upon weight status) for each phase of a program. This aggregation is shown in Figure 10. This display shows, for example, that if a program's aircraft percent weight growth is at 6% at PDR, then the program is likely, at completion, to demonstrate a cost growth of 100% (the dark blue line).

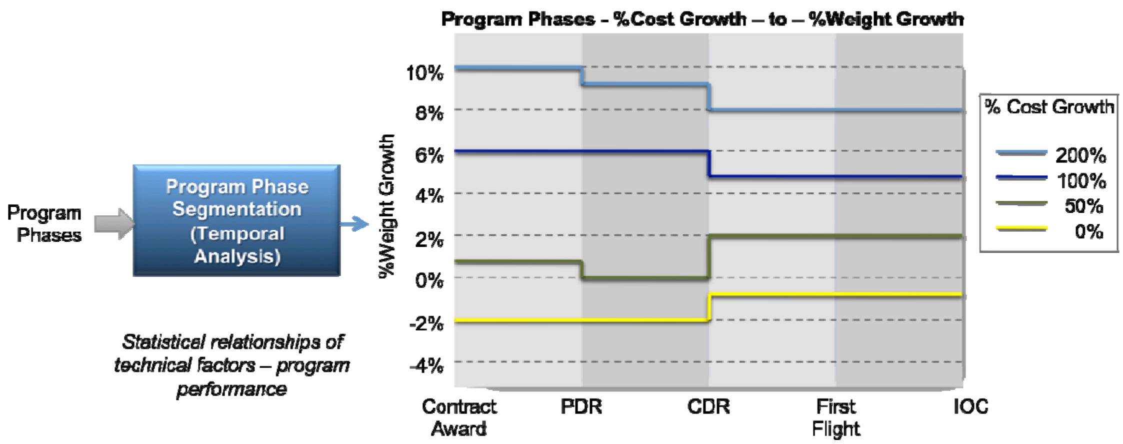


Figure 10. Segmenting ALI Statistical Analyses Into Program Phases Increases Relevance of Model

As a reminder, this data has prognostic value because it is based upon NAVAIR historical data of similar programs. As will be discussed in later sections, not all program teams welcome the analysis that their program will perform with close similarity to other programs. The reticence to accept historical coupling to their program can limit acceptance of the prognostic nature of the tool/display. User acceptance is also discussed later in the paper.

Validate fidelity and credibility of models

The correlation, regression, and time segmentation processes described previously reveal predictive strength of aircraft weight on program cost growth, especially when grouped with similar aircraft development programs and program phases. This process can, however, overaggregate the data such that a single program can overinfluence the model's predictive relevance and accuracy. To validate the model and detect such excessive influence effects, sensitivity analysis was performed on the regression analysis groups. A sensitivity analysis method called jackknife resampling was applied.

As shown in Figure 10, the regression statistics of the related family of programs (per time phase) were aggregated. Then, one-by-one, individual program data were removed from the aggregation to detect significant change in overall regression parameters. In the Figure 11 example, the blue 'X' data represent a significant departure from the aggregated regression parameters (shown with the other data markers) when a particular program was removed from the aggregation. This behavior would indicate that further examination of that excised program is necessary and presents a caution about the data as categorized and phased. If there were no departures in the data after the jackknife analysis, then confidence in the predictive regression model was increased.

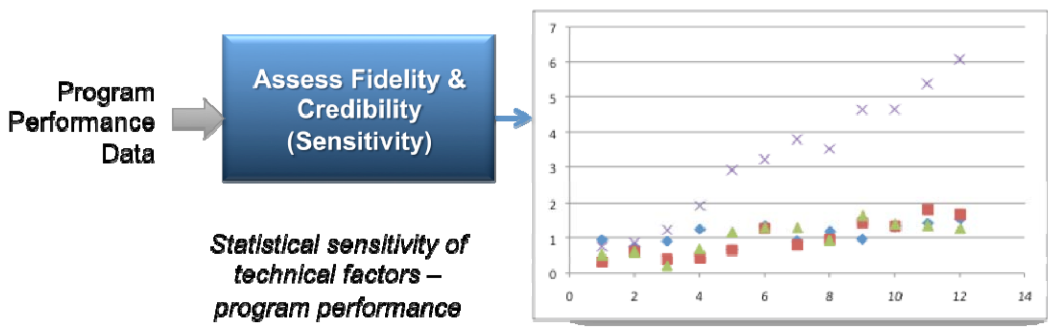


Figure 11. Statistical Resampling (Jackknife) Sensitivity Analysis Reveals Possible Dominance of a Single Program on ALI Statistical Analysis

Develop prototype tools to display system engineering leading indicators of program health based on validated models and user inputs

Throughout the statistical analysis and modeling process previously discussed, the user community (program management and engineering teams) was consulted to seed ideas about usability of an ALI tool. The previous graphical depictions were determined to be too analytic and did not have a broad appeal across all teams.

This process set out to develop a more integrated, more user-friendly ALI tool and display that integrated the statistical analysis and models previously developed. The result of this integration is shown in the primary display of the ALI tool in Figure 12. The statistical analysis results are integrated with (1) the program phases, (2) variance and uncertainty in the analysis, and (3) limits of tolerance of cost growth.

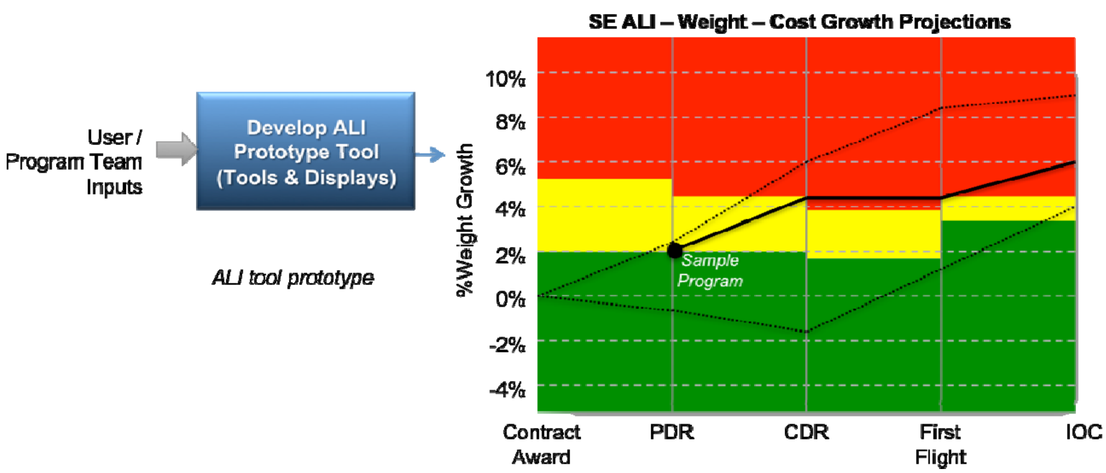


Figure 12. ALI Prototype Tool Displaying the Impact of the Current Technical Factor (e.g., Weight) Status to Projected Program Performance (e.g., Cost Growth)

In the Figure 12 example, the current status of percent weight growth is depicted by the dot. The colored bands (green, yellow, and red) are zones established by historical performance of programs that exceeded prescribed limits. These colored boundaries could be adjusted based upon the interest of the program manager, but as a minimum, would be set at cost growth conditions that would alert the program manager and leadership of severe program trouble. For NAVAIR programs, the yellow to green boundary is set at the cost growth percentage that would trigger a minor Nunn–McCurdy breach.[1] The green zone indicates the program can expect to

[1] The Nunn–McCurdy Amendment to the Defense Authorization Act of 1982 mandates that Defense-related procurement programs notify Congress when the cost of an acquisition program reaches

execute without a Nunn–McCurdy breach while a red score will likely have a Nunn–McCurdy breach. The immediate feedback provided by this type of display is an assessment of how the current program compares to previous programs and their performance related to achieving critical cost limits. In this example, the sample program dot is at the top of the green zone and could indicate that, although this program weight growth is similar to slightly “heavy” programs that went before, it may be on the cusp of “getting into trouble.”

A subtle feature of the diagram is an overall inference of the strength of its prediction based upon the data samples investigated. This strength is depicted in the size/color of the dot used to depict the current state. For example, a larger dot indicates the higher predictive strength of the underlying data. As of this writing, this feature is still being assessed for user acceptance.

The diagram provides more insight by not only assessing current status, but also by providing a sense of future performance. This prognostic feature is shown by the predictive performance line (dark black line) that predicts, based on other NAVAIR programs, that the weight of this aircraft is likely to continue to increase. The uncertainty of this prediction is depicted with the dotted confidence bounding lines (+/- 1 standard deviation range accounting for ~70% of the sample population). In this example, the program is likely to significantly exceed cost estimates (red zone) at completion. This “point estimate” based on historical data provides insight to the program leadership team to integrate into their decision-making. Such actions could include a focused weight reduction and control mitigation initiative in the development effort.

115% of the original contract amount. Additionally, if a program demonstrates a cost overrun of 25%, it will be cancelled unless the Secretary of Defense justifies its continuation to Congress.

Gather and analyze user acceptability and usability of tools

As shown above, a complex statistical analysis and model were integrated into a tool and display that provide leading indicators of a program's performance based upon engineering metrics (e.g., weight). Throughout the development process, we engaged the user community for insights into goals of the tool, usability, relevance, and areas for future growth. In many cases, the tool heightened awareness of the program teams to the usefulness of ALIs but also engendered many follow-on questions. Some examples include:

- If single-factor ALI analysis predicts cost growth, what other factors may also impact cost growth?
- What are the impact comparisons among single ALIs?
- Do other ALIs "mutual couple" to cause cost growth?
- What do I (PM/SE) do about it?
- How much is my program like historical programs?
- How can I input my own predictive performance judgment into the algorithm?

As shown in the questions above, several questions centered on multi-factor ALI impacts. The program leadership teams want to ensure that they can input current, multi-factor program metrics into the tool to provide current and high fidelity metrics into the models for incorporation into a multi-factor ALI tool.

Additionally, the models and tool are based only on historical program data related to NAVAIR ACAT I & II aircraft development programs. Feedback also indicated that the tool should ultimately be expanded to ACAT III & IV programs, subsystem upgrades, etc.

Moving to Multi-Factor ALIs

The most generalized feedback from program managers and systems engineers to the early single-factor ALI concept is that it needs (1) to consider more ALIs, (2) to incorporate their interactions, and (3) to algorithmically combine their influences into an integrated ALI metric for the program. Similar to EVM integration of cost, schedule, and achievements (milestone completion) into a few key metrics, ALI needs to work toward that goal. The process for moving to an integrated ALI output is shown in Figure 13.

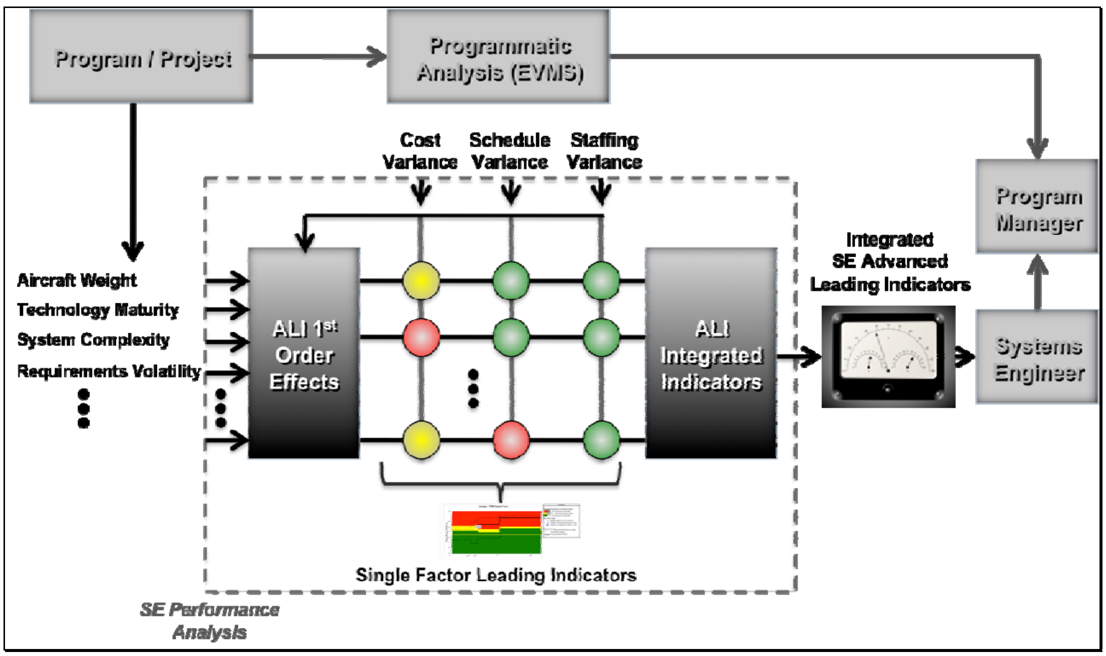


Figure 13. Single-Factor ALI Analyses Are First Steps to an Integrated ALI Output

The single-factor ALI analysis and formulations are shown in the center of the diagram. They are analyzed individually and then, after model validation, are integrated to provide a more "global" ALI metric. The repeated analysis steps are depicted in Figure 14. This process has led to an attempt at an integrated, multi-factor ALI approach that is currently being explored.

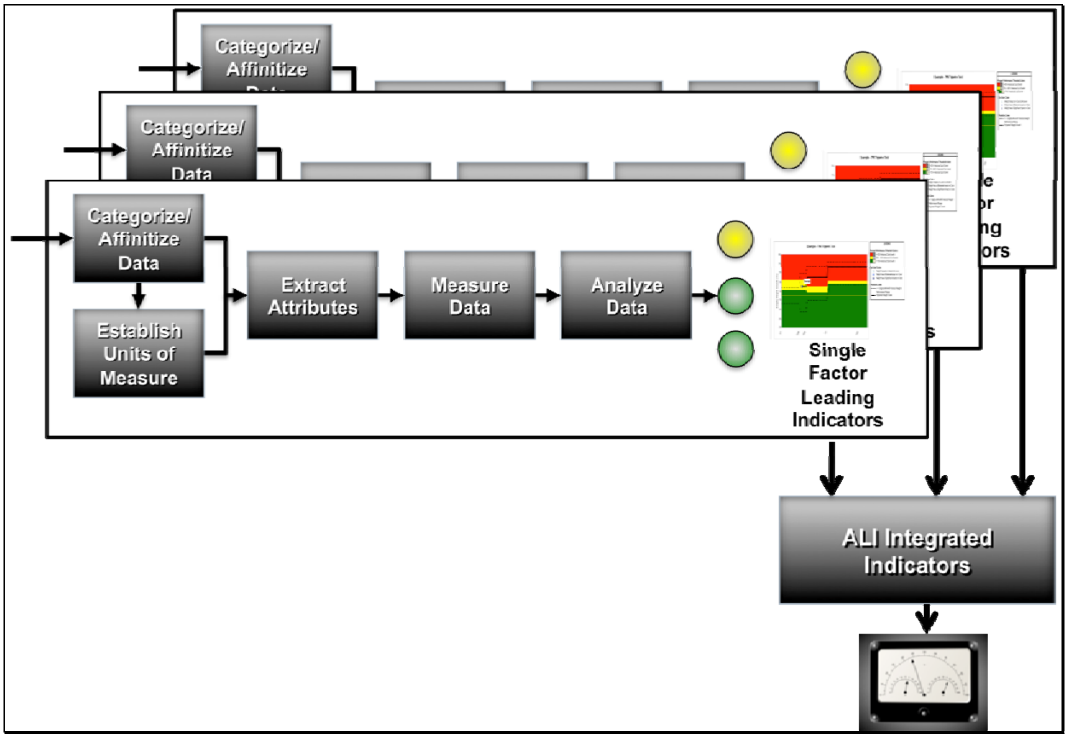


Figure 14. Parallel and Independent Single-Factor ALIs Lead to an Integrated ALI for the Program

Multi-Factor ALI Development

As discussed previously, single-factor ALI development and research has led to the current research into multi-factor ALIs. The underlying assumption is that if a single-factor ALI concept was validated historically, proved some utility in prediction program performance, and had statistical saliencies that could be exploited in a tool, then we may be able to ingest multiple ALI metrics simultaneously and provide meaningful analysis using related statistical methods suited for multi-factor analysis. Ongoing multi-factor ALI investigation does as follows (see Figure 15):

- Retains historical data analysis of key program ALI metrics. (This maintains a credible baseline of program performance upon which to compare programs.)
- Applies multiple regression methods.
- Integrates user assessment of both current conditions and their predictions of individual ALI future performance (e.g., if your program is

currently 5% over weight, what is your prediction of how this metric will change in the future?).

- Applies program end-state simulations based upon historical formulation and user estimates. After establishing both historical baseline and associated multiple regression algorithmic models, user predictions are integrated into the models via simulations to predict program performance, fit, and confidence limits.

- Provides integrated multi-factor ALI graphical output to the program leadership.

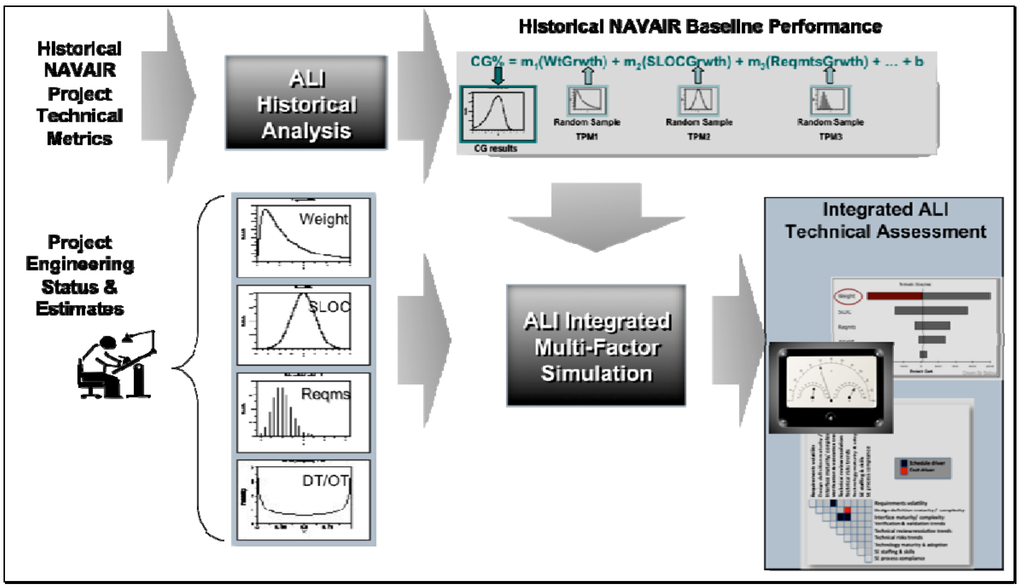


Figure 15. Multi-Factor ALI Development/Research Approach

Early graphical concepts are intended to give insights into the “mutual coupling” among the ALIs and their impact on the program. Some concepts include an “interaction matrix” approach (see the left-hand side of Figure 16) showing, for example, which multiple ALIs drive program cost and schedule (indicated by colors) and provide insight into their possible interactions (inferred by their relationships vertically and horizontally). Additionally, from multi-factor ALI analysis, it may be possible to depict which factors are most influential on program performance (see the right-hand side of Figure 16).

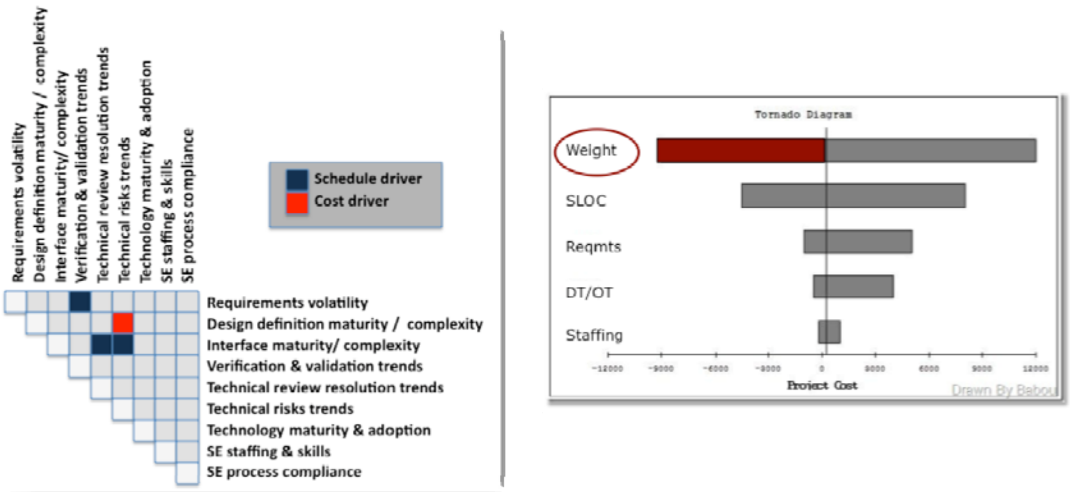


Figure 16. Example of Multiple ALIs Influencing Program Cost (Left) and Schedule and Inferring Their Possible Interactions (Vertical/Horizontal Association) and Key ALI Influencers (Right)

ALI Insight Into System Qualification Testing Success

Consistent with the authors' original goals, an NPS capstone project thesis investigated using the available ALI analysis data to gain insight into how programs were succeeding in their qualification testing (Buchanan & Jungbluth, 2010). Their research indicated some promising, although weak, statistical inferences about the data and successful testing outcomes. Their work sets foundations for further research discussed in Chapter 5.

4. Results and Conclusions

Although this ALI research is in the early stages, the ALI strategy, methods, and results discussed in this paper show promise for providing program manager and lead system engineer insight into the current and predicted technical success of their programs. This has been demonstrated through ALI data analyses, ALI user tool prototypes, and user acceptance testing.

This research began with a focus on why programs fail to meet user expectations at delivery. The goal is to determine what engineering metrics can be defined and analyzed to provide insight into success of qualification testing (e.g., operational test and evaluation, validation, etc.). This goal led us to intersect ongoing efforts related to SE ALIs that we determined would provide an understanding of closely related metrics and processes that would underpin our investigation. The ALI research is still formative and evolving and the following conclusions are mostly qualitative (non parametric) but help to refine further directions related to ALIs and the original research goals.

Data—Although there are rich data repositories available in the case of NAVAIR, the data can be inconsistent and incongruent. This increases difficulty in data analysis and bounding uncertainty in the predictive credibility of the ALI algorithms and tools. Additionally, retention of data from various programs is sometimes incomplete, leading to statistical analysis of sparse data. These problems are not, however, insurmountable and occur regularly in statistical analysis activities. The benefit of the ALI investigation is that recommended ALI metrics will emerge that can be recommended to be inculcated into the acquisitions to enable greater future ALI fidelity, granularity, and reliability.

Single-factor ALI analysis—The weight-growth versus cost-growth ALI analysis revealed that the development method was valid, provided a basis for ALI tool prototyping, and garnered preliminary user acceptance, understanding, suggested improvements, and identified ALI concept shortfalls. The technical basis is strong, however, the most impactful recommendation from users was to demand multi-factor ALI methods.

When we tried a “programmatic” metric (staffing-growth versus cost-growth) as a comparison, the statistical predictive strength was not as strong as the technical metric of weight. The resulting conclusion was that there are many external factors (rebaselining, interprogram staff balancing, etc.), which weakened statistical fit. Additionally, although we have some interest in multi-ALI interactions with programmatic metrics, we discontinued the staffing investigation because it proved too parallel with programmatic metrics (i.e., EVM).

Multi-factor analysis—These methods and analysis are in very early stages. Early models and processes are employing data from the same programs, leveraging lessons learned from single-factor analysis, expanding to include multivariate statistical methods and exploring new graphical output techniques. Early indications using simulated modeling data show promise. The next steps will include actual data, validate multivariate models, and prototype a tool to garner user acceptance.

ALI metric expansion—The only metric that was validated was aircraft weight and its growth throughout the development cycle. More metrics still need to be developed and incorporated into the research.

User acceptance—Users recognize the need for a method based upon technical metrics to provide predictive program performance insight. They do not, however, want ALI to replicate EVM-based metrics and methods. Additionally, they desire ALI methods to incorporate prediction inferences and judgments of the project engineering and management team to influence analytical output. Finally, as stated earlier, user inputs showed a strong need to reveal mutual coupling of the multiple ALI factors, the overall impact to the program, and insights into how to respond technically.

5. Areas for Continuing Research

Multi-factor ALIs—As stated previously, this analysis is in the early phases and needs to be completed to the point of testing, validation, and user acceptance/ feedback. The next steps are to include ingesting actual data, validating multivariate models, and prototyping a tool/user interface to gain insight into user acceptance

Total-Ownership-Cost control—During the conduct of this research, an acquisition emphasis change toward Total Ownership Cost (TOC) control occurred at the DoD, Department of the Navy, and NAVAIR. This potentially shifts the types of ALI metrics, but the fundamental single- and multi-factor analysis will, most likely, remain viable. The nature of a TOC data gathering, algorithm development, and tool may have to be reengineered to ensure customer acceptance and TOC problem relevance. Specifically, the following areas will need to be addressed:

- What are the salient TOC assessment goals and objectives?
- What are the ALI metrics most relevant to TOC assessment?
- What TOC ALI human interaction interfaces would be most useful to users?

Qualification and acceptance metrics—We will continue to investigate how ALI metrics (or derivatives) might be viable for also monitoring, controlling, predicting, and maximizing success of system qualification testing. A remaining goal is expanding and defining metrics and methods relative to predicting and analyzing program qualification and acceptance test success.

THIS PAGE INTENTIONALLY LEFT BLANK

Acronyms

%BNTE	Percent, below not-to-exceed
%BP	Percent, below plan
%CG	Percent, cost growth
%CWG	Percent, cumulative weight growth
ALI	Applied leading indicator
CDR	Critical design review
CTOL	Conventional takeoff and landing
EVM	Earned value management
INCOSE	International Council of Systems Engineering
KPP	Key performance parameter
KSA	Key system attribute
MOE	Measure of effectiveness
MOP	Measure of performance
NAVAIR	Naval Air Systems Command
PDR	Preliminary design review
PEO	Program Executive Office
RAM	Reliability, Availability, Maintainability
SE	Systems engineer(ing)
SEDIC	Systems Engineering Development and Implementation Center
TOC	Total ownership cost
TPM	Technical performance measure
UAS	Unmanned Aerial System
VTOL	Vertical takeoff and landing
WSARA	Weapons Systems Acquisition and Reform Act of 2009

THIS PAGE INTENTIONALLY LEFT BLANK

List of References

Arena, M. V. L., Robert S.;Murray, S. E.;Younossi, . (2006). Historical Cost Growth of Completed Weapon System Programs. *RAND Project Air Force, TR-343*.

Buchanan, C., & Jungbluth, C. (2010). *Leading Indicators for OPEVAL Success.* Unpublished capstone thesis project report, Naval Postgraduate School, Monterey, CA.

Denq, T., Hein, G., & Gilliard, J. (2009). *Applied Leading Indicator (ALI) Development Methodology, Appendix A* (working paper, NAVAIR SEDIC).

Hein, G., (2009, July). *Weight TPM Tripwire Tool Class Desk Training Package* (Unpublished PowerPoint presentation). Internal training , Naval Air Systems Command, Systems Development and Implementation Center.

Hess, R.W., & Romanoff, H.P. (1987). *Aircraft Airframe Cost Estimating Relationships: All Mission Types*, (Rand Note N-2283/1-AF). Santa Monica, CA; RAND

INCOSE. (1998). *Systems Engineering Measurement Primer (Ver. 1.0)*. Seattle, WA: INCOSE.

INCOSE. (1995). Metrics Guidebook for Integrated Systems and Product Development. Seattle, WA: INCOSE.

INCOSE. (2010). *Systems Engineering Handbook, A Guide for System LIfe Cycle Processes and Activitites v.3.2* (No. INCOSE-TP-2003-002-03.2).

Large, J.P., Campbell, H.G., & Cates, D. (1976). Parametric Equations for Estimating Aircraft Airframe Costs (R-1693-1-PA&E). Santa Monica, CA; RAND

NASA. (2007). *Systems Engineering Handbook* (NASA/SP-2007-6105 Rev 1). Retrieved from NASA website: http:www.education.ksc.nasa.gov

Rhodes, D. H., Valerdi, R., & Roedler, G. J. (2009). Systems Engineering Leading Indicators for Assessing Program and Technical Effectiveness. *Systems Engineering, 12*(1), 15.

Roedler, G. J., Rhodes, D. H., Schimmoller, H., and Jones, C. (Eds). (2010). *Systems Engineering Leading Indicators Guide, Ver 2.0*. (INCOSE -TP--2005-001-03). Seattle, WA; INCOSE.

Roedler, G. J., & Rhodes, D. H. (2007). *Systems Engineering Leading Indicators Guide, Ver 1.0*.(INCOSE -TP--2005-001-02). Seattle, WA; INCOSE.

Roedler, G. J. (2005). Technical Measurement, A Collaborative Project of PSM, INCOSE, and Industry. [INCOSE Technical Report INCOSE-TP-2003-020-01]. *INCOSE Technical Report INCOSE-TP-2003-020-01*.

Secretary of the Navy. (2008, December). *Acquisition and Capabilities Guidebook* (SECNAV M-5000.2). Washington, DC: Author

Weapons Systems Acquistion Reform Act of 2009, Public Law 111-23-May 23,2009

2003 - 2010 Sponsored Research Topics

Acquisition Management

- Acquiring Combat Capability via Public-Private Partnerships (PPPs)
- BCA: Contractor vs. Organic Growth
- Defense Industry Consolidation
- EU-US Defense Industrial Relationships
- Knowledge Value Added (KVA) + Real Options (RO) Applied to Shipyard Planning Processes
- Managing the Services Supply Chain
- MOSA Contracting Implications
- Portfolio Optimization via KVA + RO
- Private Military Sector
- Software Requirements for OA
- Spiral Development
- Strategy for Defense Acquisition Research
- The Software, Hardware Asset Reuse Enterprise (SHARE) repository

Contract Management

- Commodity Sourcing Strategies
- Contracting Government Procurement Functions
- Contractors in 21st-century Combat Zone
- Joint Contingency Contracting
- Model for Optimizing Contingency Contracting, Planning and Execution
- Navy Contract Writing Guide
- Past Performance in Source Selection
- Strategic Contingency Contracting
- Transforming DoD Contract Closeout
- USAF Energy Savings Performance Contracts
- USAF IT Commodity Council
- USMC Contingency Contracting

Financial Management

- Acquisitions via Leasing: MPS case
- Budget Scoring
- Budgeting for Capabilities-based Planning
- Capital Budgeting for the DoD
- Energy Saving Contracts/DoD Mobile Assets
- Financing DoD Budget via PPPs
- Lessons from Private Sector Capital Budgeting for DoD Acquisition Budgeting Reform
- PPPs and Government Financing
- ROI of Information Warfare Systems
- Special Termination Liability in MDAPs
- Strategic Sourcing
- Transaction Cost Economics (TCE) to Improve Cost Estimates

Human Resources

- Indefinite Reenlistment
- Individual Augmentation
- Learning Management Systems
- Moral Conduct Waivers and First-term Attrition
- Retention
- The Navy's Selective Reenlistment Bonus (SRB) Management System
- Tuition Assistance

Logistics Management

- Analysis of LAV Depot Maintenance
- Army LOG MOD
- ASDS Product Support Analysis
- Cold-chain Logistics
- Contractors Supporting Military Operations
- Diffusion/Variability on Vendor Performance Evaluation
- Evolutionary Acquisition
- Lean Six Sigma to Reduce Costs and Improve Readiness

- Naval Aviation Maintenance and Process Improvement (2)
- Optimizing CIWS Lifecycle Support (LCS)
- Outsourcing the Pearl Harbor MK-48 Intermediate Maintenance Activity
- Pallet Management System
- PBL (4)
- Privatization-NOSL/NAWCI
- RFID (6)
- Risk Analysis for Performance-based Logistics
- R-TOC AEGIS Microwave Power Tubes
- Sense-and-Respond Logistics Network
- Strategic Sourcing

Program Management

- Building Collaborative Capacity
- Business Process Reengineering (BPR) for LCS Mission Module Acquisition
- Collaborative IT Tools Leveraging Competence
- Contractor vs. Organic Support
- Knowledge, Responsibilities and Decision Rights in MDAPs
- KVA Applied to AEGIS and SSDS
- Managing the Service Supply Chain
- Measuring Uncertainty in Earned Value
- Organizational Modeling and Simulation
- Public-Private Partnership
- Terminating Your Own Program
- Utilizing Collaborative and Three-dimensional Imaging Technology

A complete listing and electronic copies of published research are available on our website: www.acquisitionresearch.org

THIS PAGE INTENTIONALLY LEFT BLANK

ACQUISITION RESEARCH PROGRAM
GRADUATE SCHOOL OF BUSINESS & PUBLIC POLICY
NAVAL POSTGRADUATE SCHOOL
555 DYER ROAD, INGERSOLL HALL
MONTEREY, CALIFORNIA 93943

www.acquisitionresearch.org